OPENING MINDS: RESEARCH FOR SUSTAINABLE DEVELOPMENT

Application of Classical Time Series Decomposition Method and Wavelet Decomposition Method to Predict the Monthly Temperature in Sri Lanka

S. Basnayake* and Ruwan Punchi-Manage

Department of Statistics and Computer Science, University of Peradeniya, Sri Lanka

**Corresponding author: Email: schamanib@yahoo.com*

1 INTRODUCTION

Joseph Fourier found that superposition of sines and cosines can be used to understand the spectral behaviour of a signal called “Fourier transformation”. However, it needs full knowledge of the signal in time domain. One of the deficiencies of the Fourier transformation is that a small alteration in the signal may affect the entire domain (Chui, 1992). In time series analysis short time Fourier transformation used a static window that moved with time. As a result, it cannot capture low frequencies and high frequencies simultaneously. Wavelet transformation is a good method to observe the data with periodic nature and sharp spikes at different scales. This method allows to capture both large scale patterns and small scale patterns at the same time by dilating the mother wavelet using a scaling vector.

In Sri Lanka there is evidence from the department of meteorology that Sri Lankan monthly average temperature has some seasonal variation. Many of them have studied the variation of temperature using classical time series decomposition method. Seasonal Autoregressive Integrative Moving Average models (*SARIMA*) is a familiar method to forecast the monthly temperature. However, only few studies used a continuous wavelet-based time series analysis to analyse the monthly temperature. Here we performed a classical time series decomposition method (*SARIMA*) and univariate continuous wavelet-based analysis for monthly temperature of Sri Lanka. We used Morelet wavelet to analyse the frequency structure of the monthly temperature in Sri Lanka. Then we compared the forecasted values and residuals from the above two methods (*SARIMA* and Wavelet decomposition) to compare the accuracy of those two models.

2 METHODOLOGY

2.1 Data

The study is based on the monthly temperature data of Sri Lanka for 115 years from 1901-2016 obtained from the World Bank climate portal (http://sdwebx.worldbank.org/climateportal). We used only 90% of the data to fit the time series models. 10% of the most recent data was used for the validation of the model.

2.2 Classical Time Series Decomposition Method

Here we used a Seasonal Auto-Regressive Integrated Moving Average model (SARIMA) to forecast the monthly temperature values in Sri Lanka. In general, SARIMA (1, 1, 1) (1, 1, 1)$_{12}$ model is given by,

$$\underbrace{(1-\phi_1 B)}_{Non-Seasonal AR(1)}\underbrace{(1-\Phi_1 B^{12})}_{Seasonal\ AR(1)}(\underbrace{1-B)}_{\substack{Non-seasonal\\ difference}}\underbrace{(1-B^{12})}_{\substack{Seasonal\\ difference}}y_t = \underbrace{(1+\theta_1 B)}_{Non-Seasonal MA(1)}\underbrace{(1+\Theta B^{12})}_{Seasonal\ MA(1)}$$

In classical time series approach, first the data set was decomposed and the seasonality and trend components were identified. Since the data set exhibited a seasonality component we used Seasonal Auto Regressive Integrated Moving Average (SARIMA) model. The best fitting parameters were identified using the auto.arima function in R based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values. The residual analysis was carried out to check whether the model performed well with the data. In the second step we tested whether residuals are normally distributed around zero mean and constant variance. Box-pierce test is performed to test whether residuals follows a white noise series with zero mean. Then the SARIMA model with the best fitting parameters was used to forecast the temperature for the next 12 years. Actual and fitted value plots were used to see how well the model behaves for the forecasted period. All analysis were performed using the Package “forecast” and “tseries” in (R core team, 2016).

2.3 Wavelet Method:

In the mother Morelet wavelet is given by,

$$\psi(t) = \pi^{-1/4} e^{i\omega t} e^{-\frac{t^2}{2}}.$$

The Morelet wavelet transform of a time series (x_t) is defined as the convolution of the series with a set of “wavelet daughters” generated by the mother wavelet by transition in time by τ and scaling by s

$$Wave(\tau, s) = \sum_t x_t \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-\tau}{s} \right) .).$$

The position of the particular wavelet is determined by τ (localizing time parameter) and s is the wavelet coverage in the frequency domain. S has a number of octaves (mother Wavelet is double with each octave) and number of voices per octaves (see details for Roesch and Schmidbauer, 2014).

ISSN 2012-9916 © The Open University of Sri Lanka

3 RESULTS AND DISCUSSION

According to classical time series decomposition method a clear trend is observed (tr_t) (Figure 1a). A seasonal fluctuation (S_t) is observed with every 12 time period (Figure 1a).

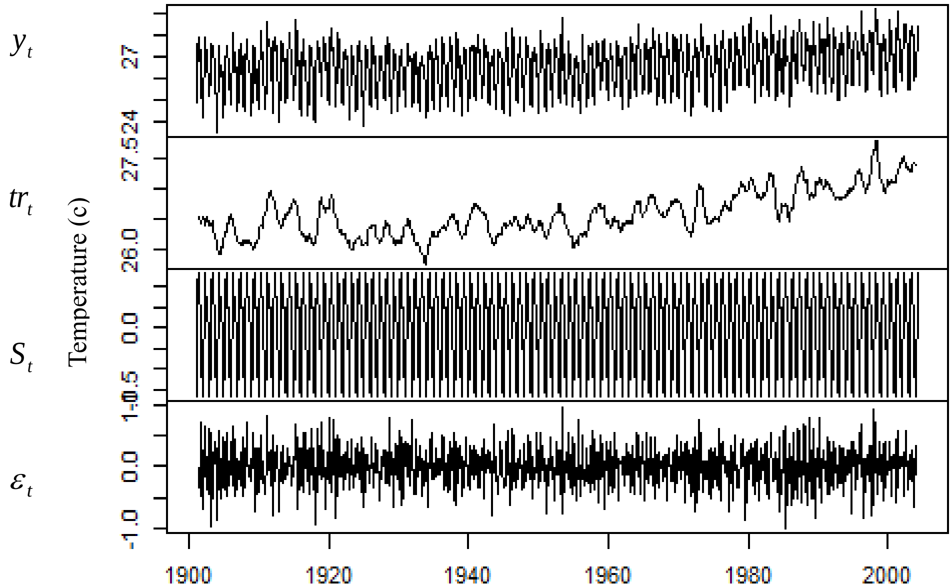


Figure 1a: Time series decomposition using R

Residuals fluctuate around a constant zero mean and constant variance (Fig. 1a). We found that the best model is SARIMA (5, 1, 0) (1, 1, 0)$_{12}$. Box pierce test statistic (Chi. Sq. value 0.084987, df = 1) gives the p value of 0.7706. This indicates that residuals follow a white noise series with a zero mean and constant variance.

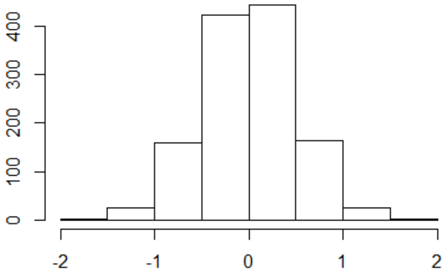


Figure 1b. Histogram for the errors

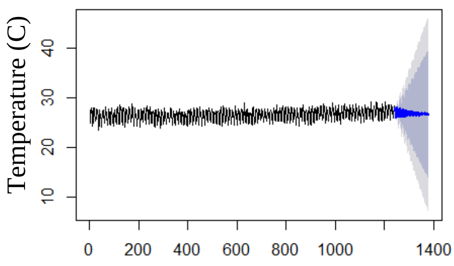


Figure 1c. Forecasting

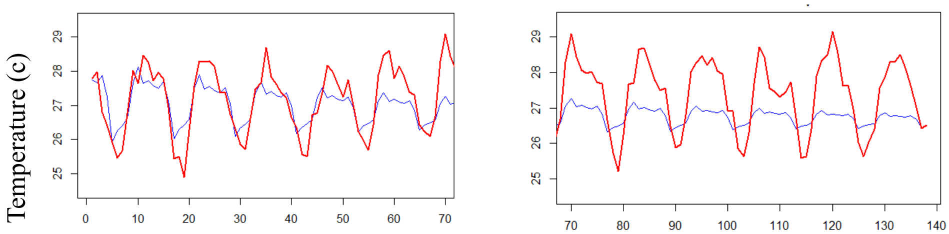


Figure 1d: Fitted values (in blue) and observed values (in red).

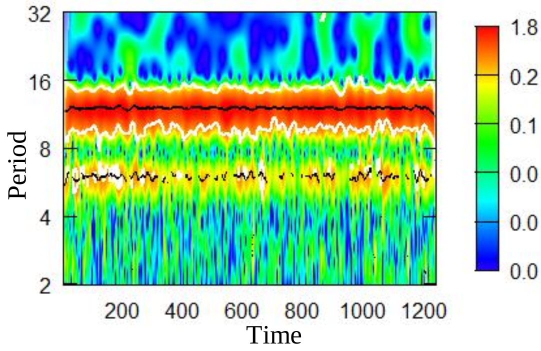
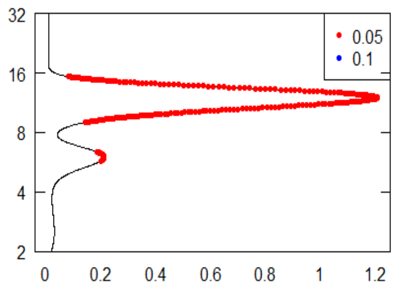


Figure 2a: Wavelet decomposition

(Power spectrum of the series, method= “white noise”)



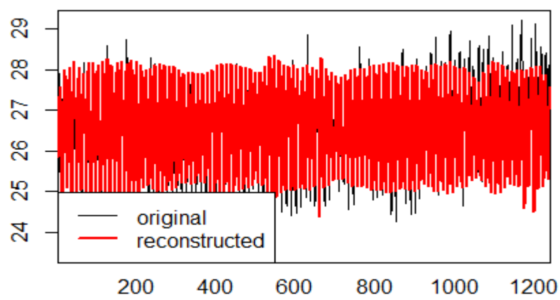


Figure 2b: Average power series **Figure 2c:** Reconstruction of time series using wavelet.

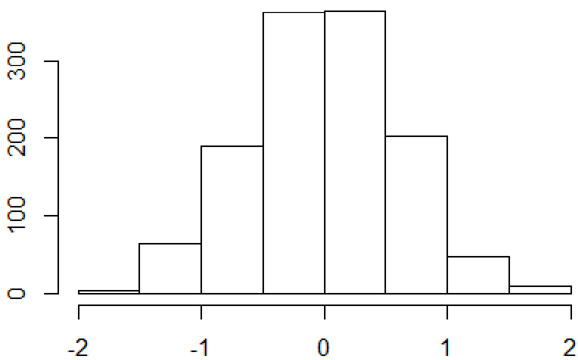


Figure 2d: Histograms of errors from wavelet decomposition meth

ISSN 2012-9916 © The Open University of Sri Lanka

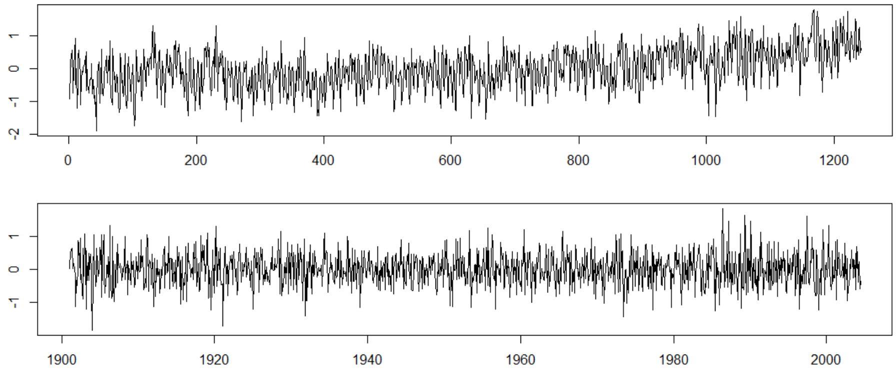


Figure 2e: Upper panel: residuals from classical time series decompositon. Lower Panel: residuals from wavelet decomposition method.

According to the wavelet method we found that the average power series peaked at the 12^{th} period. Seasonality at 12^{th} period is expected (Fig. 2b). Because in classical best fitted SARIMA model s=12. However, we noticed that the average power series peaked at 6^{th} period as well. We hadn't observed this from the SARIMA models. The reconstructed time series using the wavelet method is showed in Fig 2c. We test the residuals obtained from the wavelet method to see whether they distributed with a zero mean and constant variance (Fig 2d). Box-pierce test for residuals gives p value of 0.08148 which do not reject the null hypothesis that residuals follow a white noise series.

4 CONCLUSIONS AND RECOMMENDATIONS

In classical time series analysis seasonality was tested by auto correlation (ACF) and partial auto correlation function (PACF). However, wavelet transformation outperforms classical time series decomposition when data contains two or more seasonal super-positions of ~~a~~ data. This indicates that wavelet method is useful alternatively to understand multiple seasonality of a time series.

REFERENCES

A. Trapletti and K.Hornik (2017). tseries: Time Series Analysis and Computational Finance. R package version 0.10-40.

A.Roesch and H. Schmidbauer (2014). WaveletComp: Computational Wavelet Analysis. R package version 1.0.

C.K. Chui (1992) *An Introduction to Wavelets, Academic press, London, UK*

Data:http://sdwebx.worldbank.org/climateportal/index.cfm?page=downscaled_data_download and menu=historical

Hyndman R.J. and Khandakar, Y. (2008). "Automatic time series forecasting: the forecast package for R." _Journal of Statistical Software, *26*(3), pp. 1-22

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.